

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

Technical Report 32-1165

Revision 1

*Reliability-Confidence Combinations for Small-Sample
Tests of Aerospace Ordnance Items*

A. G. Benedict

FACILITY FORM 602

N 68-28775 (ACCESSION NUMBER)

17 (PAGES)

ck-95397 (NASA CR OR TMX OR AD NUMBER)

(THRU)

1 (CODE)

15 (CATEGORY)

GPO PRICE \$ _____

CSFTI PRICE(S) \$ _____

Hard copy (HC) _____

Microfiche (MF) _____

ff 653 July 65

JET PROPULSION LABORATORY
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA

July 1, 1968



NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

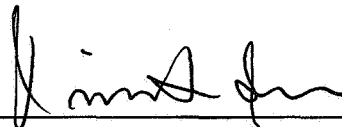
Technical Report 32-1165

Revision 1

*Reliability-Confidence Combinations for Small-Sample
Tests of Aerospace Ordnance Items*

A. G. Benedict

Approved by:

A handwritten signature in black ink, appearing to read 'Winston Gin', is written over a horizontal line.

Winston Gin, Manager
Solid Propellant Engineering Section

JET PROPULSION LABORATORY
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA

July 1, 1968

TECHNICAL REPORT 32-1165

Copyright © 1968
Jet Propulsion Laboratory
California Institute of Technology
Prepared Under Contract No. NAS 7-100
National Aeronautics & Space Administration

PRECEDING PAGE BLANK NOT FILMED.

Acknowledgment

Appendix B of this paper was prepared by R. G. Chamberlain of JPL whose contribution here and whose constructive suggestions elsewhere in preparation of this revision are gratefully acknowledged.

Contents

I. Introduction	1
II. Inherent Reliability of 'No-History' Lots	2
III. Confidence and Unexpected Fail Rates	3
IV. Reliability-Confidence Combinations	3
V. Compounding Subsystem Reliabilities	4
VI. Tests by Attributes vs Tests by Variables	5
VII. Conclusions	6
Appendix A. Reliability of Simple Series and Parallel Systems Based on a No-Fail, Small-Sample Test of the Lot From Which the Components Are Drawn	7
Appendix B. Sifting Effect of Small-Sample, No-Fail Tests	9

Tables

1. Some reliability-confidence combinations for a five-sample, no-fail test	2
2. Approximate chances of various fail rates for a lot yielding no fails in eight samples	3
3. Effect of using component reliabilities corresponding to $\gamma = 70\%$ for calculation of reliability of systems compounded of items drawn from the one lot	5

Figures

B-1. Number of lots of each reliability after n no-fail tests	10
B-2. Sifting effect on high-reliability lots	11

Abstract

Reliability-confidence combinations for small-sample, no-fail tests of aerospace ordnance items are considered in some detail. Analyses epitomized by the widely used, but sometimes misapplied, equation

$$\gamma = 1 - R^n$$

are shown to provide unexpectedly good approximations. Component reliabilities corresponding to confidences of the order of 70%, as based on tests by attributes, are shown to be satisfactory for calculation of simple series and parallel system reliabilities; it is suggested that 70% confidence levels may also be a good choice for tests by variables.

Reliability-Confidence Combinations for Small-Sample Tests of Aerospace Ordnance Items

I. Introduction

For a complex aerospace system to have even a 50% chance of performing its intended function, the chances of failure in any one critical subsystem component must be extremely low.

In the evaluation of the reliability of aerospace ordnance components, it is common to find that

- (1) The items are expensive, limiting the number of samples available.
- (2) The items are of a special design, with no history to suggest their reliability.
- (3) Because the target reliability approaches unity, evaluation should be directed at fail rates rather than reliability.¹
- (4) The items to be tested are "single-shot" and, unlike solenoid relays for example, cannot be operated nondestructively; items operated in test cannot be used in flight.

¹A decrease of less than 1% in reliability, from 99.9% to 99%, would represent a tenfold increase in fail rate (i.e., from $1/1000$ to $1/100$).

- (5) Tests by both attributes and variables may be appropriate, but the frequency distribution for tests by variables is unknown.²
- (6) There is some uncertainty (if not confusion) in the selection of a reliability-confidence combination that best expresses the results of the evaluation.

Today, evaluation of small-sample aerospace ordnance tests follows along lines developed in the 1920s by Western Electric and Bell Telephone engineers³ as characterized by the equation

$$\sum_{j=0}^{j=P} \binom{n}{j} f^j (1-f)^{n-j} = \sum_{j=0}^{j=P} \binom{n}{j} (1-R)^j R^{n-j} = (1-\gamma) \quad (1)$$

²Log-normal or comparable distributions are often assumed simply as a convenient economy. Although actual test results involving only small samples may prove to be consistent with some arbitrary distribution, use of such a preselected distribution for extrapolation to extreme percentiles may be grossly misleading.

³Dodge, H. F., and Romig, H. G., *Sampling Inspection Tables, Single and Double Sampling*, Second Edition, p. 1. John Wiley & Sons, Inc., New York, 1959.

where

n = number of pass/fail tests

F = number of fails detected

$$\binom{n}{f} = \frac{n!}{(n-f)!f!}$$

R = assumed reliability

f = assumed fail rate

γ = lower confidence limit on R
(or upper confidence limit on f)

Equation (1) reduces, for a series of n tests with no failures,⁴ to

$$(1-f)^n = R^n = 1-\gamma \quad (2)$$

Both equations allow the results of any particular test program to be expressed as indicating any of an infinite number of reliability-confidence combinations. For example, a five-sample, no-fail test could be interpreted as indicating any one of as many combinations as one wished, including those shown in Table 1, as calculated from Eq. (2).

Table 1. Some reliability-confidence combinations for a five-sample, no-fail test^a

Reliability R , %	Confidence γ , %	Maximum fail rate f , %
99	5	1
90	41	10
87	50	13
80	67	20
70	83	30
60	92	40
50	97	50
40	99	60

^aNote the rapid rise in confidence for a decrease in reliability from 99% to 90%. It may be shown that, for $f \approx 0$, $\log_{10}(1-\gamma) \approx -0.434 nf$.

The question naturally arises as to which combination, if any, is the most descriptive. For example, if the five-sample test were used to judge the quality of a further six items drawn from the same lot, would the 80% reliability-67% confidence combination imply that, of the six samples, at least five (80%), or four (67%), or three (80% of 67%),

⁴For simplicity, subsequent discussion is limited to no-fail tests; these are of particular interest in aerospace ordnance because the number of samples available for test is usually so small that even one failure may imply an intolerably high fail rate.

should be expected to pass, or should none be expected to pass because none are 100% reliable?⁵

Before proceeding, it is important to note the original problem faced by the Western Electric and Bell Telephone engineers was how to use small samples to determine whether a particular incoming shipment of components met a quality level which prior shipments had shown to be practical. If earlier shipments had exhibited a reliability of 98% or better, with an occasional bad lot having a reliability as low, for example, as 80%, the customer might be content with a small-sample receiving inspection scheme that gave him a good chance of detecting lots with a reliability of less than 98%. By comparison, many aerospace ordnance items involve short-run, one-time production offering no prior history on which to base an expected reliability; this raises the further question of whether Eq. (2) is grossly inappropriate in such cases.

II. Inherent Reliability of 'No-History' Lots

At first glance, it might seem that evaluation of a lot for which all fail rates between 0% and 100% were equally likely would be a much more pessimistic process than evaluation of a lot for which fail rates below, say, 70% were assumed to be somewhat unlikely. This is not so, however, because small-sample, no-fail tests quickly cull (or "screen") high fail-rate lots (see Appendix B).

If an n -sample test is made on a large number K of lots of size L for which all proportions p of passes P between 0 and 1 are equally likely, the chances C of any lot yielding a sample of n P 's is given by⁶

$$C = p^n \quad (3)$$

and the number of lots M that will yield such samples is given by

$$M = \int_0^1 KC dp = \int_0^1 Kp^n dp$$

⁵Although Eqs. (1) and (2) relate to tests by attributes, reliability-confidence combinations are also used to express the outcome of tests by variables, giving rise to the same question.

⁶Equation (3) is true only if samples are returned to the lot as drawn, or if the sample size n is so small by comparison with the lot size L that the removal of the sample has no significant effect on the proportion p of P 's remaining.

from which

$$M = \frac{K}{n+1} \quad (4)$$

The proportion R_i of P 's in these M lots representing their average quality or "inherent reliability" is given by

$$R_i = \frac{1}{ML} \int_0^1 KLp^n p dp = \frac{KL}{ML} \times \frac{1}{n+2} = \frac{n+1}{n+2} \quad (5)$$

and the proportion γ of the M lots that will have at least any proportion R of P 's is given by

$$\gamma = (n+1) \int_R^1 p^n dp$$

from which

$$\gamma = 1 - R^{n+1} \quad (6)$$

Note the similarity between Eq. (2) and Eq. (6). Equation (6) shows that an isolated lot yielding no fails in a sample of size n will have (with confidence γ) a comparatively high reliability even if, prior to sampling, all reliabilities between 0 and 1 were considered equally likely. It can be shown that although the inherent reliability approaches 1 as n approaches ∞ , the corresponding γ converges on only $1 - e^{-1\infty}$ ($= 63.22\%$).

Although Eq. (2) would not be strictly applicable to the case under consideration, its misuse would still give a good approximation, i.e.,

$$\gamma = 1 - R^{n+1} \approx 1 - R^n$$

III. Confidence and Unexpected Fail Rates

Although a proportion $\gamma = 1 - R^{n+1}$ of lots yielding no fails in samples of size n will contain a proportion $(1 - R) = 1/(n+2)$ or less of faulty items, the remaining fraction R^{n+1} of the lots will contain a proportion of faulty items larger than $1/(n+2)$; from these lots will come fail rates which may be unexpected and possibly disappointing. For example, from an eight-sample, no-fail draw, the probability of various fail rates can be illustrated as in Table 2.

Table 2. Approximate chances of various fail rates for a lot yielding no fails in eight samples

Reliability R^a	Confidence γ^b	Most likely maximum fail rate ($1 - R$)	Approximate chances of indicated fail rate ^c
1	0	0	0
0.9	0.613	0.1	$(0.613-0) = 0.61$
0.8	0.864	0.2	$(0.864-0.613) = 0.25$
0.7	0.959	0.3	$(0.959-0.864) = 0.10$
0.6	0.9897	0.4	$(0.9897-0.959) = 0.03$
0.5	0.9980	0.5	$(0.9980-0.9897) = 0.008$
0.4	0.9997	0.6	$(0.9997-0.9980) = 0.002$

^aThe inherent reliability of Eq. (5) is 0.9.
^bCalculated from Eq. (6).
^cThese chances relate strictly to ranges of fail rates, but are more easily visualized as relating to fails in a lot of size 10.

From Table 2, an expectation that the fail rate would be 1/10 or less would result in a 25% chance of being disappointed by a factor of 2, and a 10% chance of being disappointed by a factor of 3. In a manufacturing process (such as assembly of radio receivers), inadvertent acceptance of one lot in four of components (such as capacitors) with a fail rate twice "normal," or of one lot in ten with a fail rate three times "normal," might be countered by replacement of faulty items when the assembly failed to meet final tests: low confidence in component reliability might not necessarily be reflected in the reliability of the manufactured system, but would almost certainly be reflected in the cost of "rework" during manufacture. In such a case, the selection of the optimum reliability-confidence combination would involve trade-offs between the costs of additional acceptance testing, on the one hand, with the costs of rework, on the other.

In aerospace ordnance systems, however, faulty sub-components may not be detectable by "final inspection" of a system, and low confidence in component reliability may consequently manifest itself in low system reliability.

IV. Reliability-Confidence Combinations

For a simple series system (such as a chain of links) where N components each of pass rate p are used, the system reliability R_s is a probability given by

$$R_s = p^N \quad (7)$$

Similarly, for a simple parallel system (such as a raft made from oil drums) where N components of pass rate p

are used and inclusion of one faulty item is tolerable, the system reliability R_p is a probability given by

$$\begin{aligned} R_p &= p^N + N p^{N-1} (1 - p) \\ &= N p^{N-1} - (N - 1) p^N \end{aligned} \quad (8)$$

When the reliability R but not the true pass rate p of components in series or parallel systems is known, it is common practice to substitute R for p in equations such as Eqs. (7) and (8), often using component reliabilities calculated to an arbitrary confidence. For any given system, the system reliability so calculated will obviously depend on the confidence associated with the component reliabilities; it is system reliabilities qualified by confidence that are often difficult to interpret.

Although small-sample tests do not allow determination of the true pass rate p , in Appendix A it is shown that a no-fail test of n samples allows calculation of the system reliabilities of Eqs. (7) and (8) as follows:

$$R_s = \frac{n + 1}{N + n + 1} \quad (9)$$

$$R_p = \frac{n + 1}{N + n} \times \frac{2N + n}{N + n + 1} \quad (10)$$

Even though no-fail tests by attributes thus allow calculation of the probability of success of a system without expressing the quality of the components in the form of a reliability-confidence combination, it may nevertheless be of interest to consider the practicality of substituting some such reliability for p in Eqs. (7) and (8) to yield useful approximations for R_s and R_p as calculated by Eqs. (9) and (10). In essence, the problem is to find some value of R which will satisfy the relationships

$$R_s = R^N = \frac{n + 1}{N + n + 1} \quad (11)$$

$$R_p = N R^{N-1} - (N - 1) R^N = \frac{n + 1}{N + n} \times \frac{2N + n}{N + n + 1} \quad (12)$$

Any such value of R should be substantially independent of N to cover the common situation in which the nature of the system is unknown by those conducting the small-sample test.

In Appendix A, it is also shown that R corresponding to a confidence of about 63.2% can provide a useful approximation if substituted in Eq. (11), and that R corresponding to a confidence of about 75.7% can provide a useful approximation if substituted in Eq. (12). However, the component might be used in either a series or a parallel system; Table 3 illustrates that use of reliabilities calculated to a 70% confidence level (as a compromise between 63.2% and 75.7%) provides reasonable approximations not only for series systems, but also for parallel systems, at least over a wide range of practical situations. This table relates to tests by attributes; although the comparable situation for tests by variables has not been explored, it seems likely that component reliabilities expressed at a 70% confidence level as the result of tests by variables might also compound to yield fair approximations for certain series and parallel systems.

On the other hand, Table 3 shows that system reliabilities calculated from component reliabilities based on small-sample tests are, at best, only approximations; if the true (most likely) pass rate of a system can be calculated from no-fail tests by attributes, why then resort to expressing component reliabilities at some arbitrary confidence level? The answer to this question lies in the following:

- (1) All calculated pass rates are subject to error arising from the assumed distribution. The equations of this report are based on the assumption that the lot selected for sampling and use is one of an infinite number of lots for which all fail rates are equally likely; if another distribution prevailed, the discrepancies of Table 3 might prove relatively insignificant.
- (2) The results of small-sample tests by variables are probably most easily expressed in terms of reliability demonstrated at a particular confidence level; while this situation prevails it seems undesirable to use a different form of expression for tests by attributes.

V. Compounding Subsystem Reliabilities

It has been shown that, under certain circumstances, the reliability of a system can be expressed as (or approximated by) a true pass rate; if such a system is a (series) subsystem in a larger system for which no subsystem failures are tolerable, the reliability (or pass rate) of the whole system R_t will be the product of the reliabilities (pass rates) of the various subsystems R . In the simple

Table 3. Effect of using component reliabilities corresponding to $\gamma = 70\%$ for calculation of reliability of systems compounded of items drawn from the one lot^a

Number of samples in no-fail test (n)	Reliability R at 70 % confidence, $= (1 - 0.70)^{1/(n+1)}$ $= 0.30^{1/(n+1)}$	Series systems, N components, no faulty components tolerable						Parallel systems, N components, not more than one faulty component tolerable			
		True system pass rate: $R_s = \frac{n+1}{N+n+1}$ Pass rate: $f_1(R) = R^N$						True system pass rate: $R_p = \frac{n+1}{N+n} \times \frac{2N+n}{N+n+1}$ Pass rate: $f_2(R) = NR^{N-1} - (N-1)R^N$			
		N = 2		N = 10		N = 100		N = 2		N = 100	
		$1 - R_s$	$1 - f_1(R)$	$1 - R_s$	$1 - f_1(R)$	$1 - R_s$	$1 - f_1(R)$	$1 - R_p$	$1 - f_2(R)$	$1 - R_p$	$1 - f_2(R)$
2	0.669	0.40	0.55					0.100	0.109		
3	0.740	0.33	0.45					0.067	0.068	0.49	0.78
4	0.786	0.29	0.38					0.048	0.046	0.43	0.66
8	0.875	0.18	0.23					0.018	0.016	0.26	0.36
10	0.896	0.15	0.20	0.48	0.67			0.013	0.011	0.21	0.28
20	0.9443	0.087	0.108	0.32	0.44			0.0040	0.0031	0.097	0.104
40	0.9711	0.047	0.057	0.20	0.25			0.0011	0.00084	0.035	0.032
80	0.9852	0.024	0.029	0.11	0.14			0.00029	0.00022	0.011	0.0091
100	0.9882	0.019	0.024	0.090	0.112	0.50	0.70	0.00019	0.00014	0.0074	0.0059
200	0.9940	0.0099	0.012	0.047	0.058	0.33	0.45	0.000049	0.000036	0.0020	0.0016
400	0.9970	0.0050	0.0060	0.024	0.030	0.20	0.26			0.00053	0.00040
1000	0.99880	0.0020	0.0024	0.0099	0.012	0.091	0.11			0.000088	0.000065
2000	0.99940	0.0010	0.0012	0.0050	0.0060	0.048	0.058				
4000	0.99970	0.00050	0.00060	0.0025	0.0030	0.024	0.030				
10000	0.99988	0.00020	0.00024	0.0010	0.0012	0.0099	0.012				

^aCalculated failure rates $(1 - R)$, rather than calculated reliabilities (R) have been tabulated for ease of comparison. Thus, from the table, for a series system of 10 units compounded from a lot which had yielded no fails in a sample of 20, $1 - R^{10}$ with R calculated for $\gamma = 70\%$ would yield a maximum fail rate of 0.44, 37% $\left(\text{i.e., } \frac{0.44-0.32}{0.32} \right)$ higher than the likely (true) fail rate.

case where n subsystems of equal reliability R are involved,

$$R_t = R^n$$

Although R_t and R are both pure pass rates without a qualifying confidence level, the qualities of both the system and subsystems are possibly easier to understand as confidences (in 100% reliability) than as reliabilities (at 100% confidence).

VI. Tests by Attributes vs Tests by Variables

Present-day qualification tests for aerospace ordnance items are, more often than not, based on tests by attributes—a carry-over from typical qualification tests for military ordnance items. Unlike aerospace ordnance items, many military items involve long production runs with fixed tooling, are not likely to be used as critical elements in a complex and expensive system, and can be recalled if intolerable failure rates are detected in service; although qualification to a reliability of possibly 95% might be acceptable for a military cartridge, the different cir-

cumstances might make qualification to a reliability of 99.9% more appropriate for the aerospace counterpart of this same cartridge.

Qualification by attribute-testing to a 99.9% reliability requirement (at a 70% confidence level) would call for a minimum of about 1200 samples for each type of destructive test—a quantity which would be economically prohibitive in many cases.

Provided one is prepared to assume an arbitrary distribution (such as the log-normal) and can devise appropriate tests, qualification by use of tests by variables may, by contrast, allow use of a practical minimum of only about six samples.

Although use of tests by variables is consequently attractive (if not unavoidable) for qualification of aerospace ordnance items, until the problem of reliability-confidence combinations for such tests has been solved, the only practical course seems to be to adopt the 70% confidence level which appears appropriate for tests by attributes.

VII. Conclusions

If, prior to testing, it is agreed that the lot of items is one of many for which all fail rates between 0 and 1 are equally likely

- (1) Analysis of test results by methods epitomized by the equation

$$\gamma = 1 - R^n$$

can continue to be used without introducing serious errors.

- (2) Samples of size n that yield no failures belong to families of lots having an inherent reliability

$$R_i = \frac{n+1}{n+2}$$

with a confidence $\gamma = 1 - R^{n+1}$.

- (3) The true pass rate of a simple series or parallel system or subsystem can be calculated on the basis

of the results of no-fail tests by attributes, without reference to a confidence, but is possibly easiest understood if described as a confidence (in 100% reliability).

- (4) Component reliabilities expressed at a 70% confidence level as the result of no-fail tests by attributes may be compounded to yield fair approximations to pass rates for simple series and parallel systems, at least over a wide range of practical situations. It seems reasonable to assume that the results of tests by variables expressed as reliabilities at a confidence of 70% could also be compounded to yield fair approximations to the pass rates of such systems.
- (5) Although confidence levels of about 70% appear appropriate in calculation of system or subsystem reliabilities, higher or lower confidence levels may be more appropriate from some economic standpoint.

Appendix A

Reliability of Simple Series and Parallel Systems Based on a No-Fail Small-Sample Test of the Lot From Which the Components Are Drawn

Let

R_s = reliability of a simple series system (e.g., a chain of links) in which no faulty components are tolerable

R_p = reliability of a simple parallel system (e.g., a raft of empty oil drums) in which not more than one faulty component is tolerable

n = number of samples (chain links or oil drums) in a no-fail test of the lot from which the system is compounded

N = number of components in the system

p = true pass rate of the components

R_i = inherent reliability of the components $(n + 1)/(n + 2)$

$\gamma = 1 - R^{n+1}$ = Confidence in component reliability R

Assume that, prior to sampling, the lot of components from which the system is to be compounded is one of a large number of lots for which all pass rates p between zero and one are equally likely.

I. Series Systems

The proportion of lots which will yield n good samples in an n -sample test = $1/(n + 1)$, and of lots which will yield $n + N$ good samples in an $(n + N)$ draw = $1/(N + n + 1)$ (see Eq. 4).

The chances of getting N good samples in an N -sample draw after having previously drawn no fails in an n -sample draw is, consequently, given by

$$\frac{n + 1}{N + n + 1}$$

$$\therefore R_s = \frac{n + 1}{N + n + 1}$$

But

$$R_s = p^N$$

$$\therefore p = \left(\frac{n + 1}{N + n + 1} \right)^{1/N} \approx \left(1 - \frac{N}{n + 1} \right)^{1/N} \quad N \ll n$$

$$\approx 1 - \frac{1}{n + 1} \approx \frac{n + 1}{n + 2} \quad n \gg 1$$

$$\therefore p \approx R_i$$

Thus, for series systems in which $N \gg n$ and $n \gg 1$, the inherent reliability of components may be used as a true (unqualified) pass rate for calculation of the reliability of the system; alternatively, because the confidence for inherent reliability converges on 63.2%,⁷ the reliability corresponding to a confidence of 63.2% could be used as the true pass rate.

II. Parallel Systems

The fraction of lots which will yield n good samples in n is given by

$$\int_0^1 p^n dp = \frac{1}{n + 1} = F_1 \text{ (see Eq. 4)}$$

The fraction of lots which will yield no failures in a sample of size $(N + n)$ is similarly given by $1/(N + n + 1) = F_2$

⁷Even for small values of n , the reliability at a confidence of 63.2% does not differ markedly from the inherent reliability, nor does the confidence in inherent reliability differ markedly from 63.2%. These points are shown by the following:

n	2	4	7	15	80
$R_i = \frac{n + 1}{n + 2}$	0.75	0.833	0.889	0.941	0.988
R for 63.2% confidence, $= (1 - 0.632)^{1/(n + 1)}$	0.717	0.819	0.883	0.939	0.988
Confidence in inherent reliability, $\gamma = 1 - R_i^{n+1}$	0.58	0.60	0.61	0.62	0.63

and the fraction of lots which will yield no more than one failure is given by

$$\begin{aligned} & \int_0^1 [p^{(N+n)} + (N+n)p^{(N+n+1)}(1-p)] dp \\ &= \int_0^1 [(N+n)p^{(N+n-1)} - (N+n+1)p^{(N+n)}] dp \\ &= \frac{2}{N+n+1} = F_3 \end{aligned}$$

Thus, a fraction $F_3 - F_2 = 1/(N+n+1) = F_4$ of lots will yield one failure in $(N+n)$ samples, and of this fraction a proportion $N/(N+n) = F_5$ involves no failures in the first n samples.

The proportion of lots R_p which, having yielded n no-fail samples, will then yield no more than one failure in the next N samples is, therefore, given by

$$\begin{aligned} R_p &= \frac{F_2 + F_5 \times F_4}{F_1} \\ &= \frac{\frac{1}{N+n+1} + \frac{N}{N+n} \times \frac{1}{N+n+1}}{\frac{1}{n+1}} \\ &= \frac{n+1}{N+n} \times \frac{2N+n}{N+n+1} \end{aligned}$$

But

$$\begin{aligned} R_p &= p^N + Np^{N-1}(1-p) \\ &= Np^{N-1} - (N-1)p^N \\ &= N(1-f)^{N-1} - (N-1)(1-f)^N \quad f = 1-p \\ &= 1 - \frac{N(N-1)}{2!} f^2 \quad Nf \ll 1 \\ &= \frac{n+1}{N+n} \times \frac{2N+n}{N+n+1} \\ \therefore f &= \left[\frac{2}{(N+n)(N+n+1)} \right]^{1/2} \quad Nf \ll 1 \end{aligned}$$

Now

$$\gamma = 1 - R^{n+1}$$

and for

$$R = p$$

$$\gamma = 1 - p^{n+1} = 1 - (1-f)^{n+1}$$

$$\begin{aligned} &= 1 - \left[1 - \left(\frac{2}{(N+n)(N+n+1)} \right)^{1/2} \right]^{n+1} \\ &= (n+1) \left[\frac{2}{(N+n)(N+n+1)} \right]^{1/2} - \frac{(n+1)(n)}{2!} \\ &\quad \times \left[\frac{2}{(N+n)(N+n+1)} \right] + \dots \end{aligned}$$

Which, for $n \gg N$

$$\approx 2^{1/2} - \frac{2}{2!} + \frac{2}{3!} - \dots$$

i.e.,

$$\gamma \approx 1 - e^{-\sqrt{2}} = 1 - \frac{1}{4.113} = 75.7\%$$

Thus, for parallel systems in which one faulty component is tolerable and the number of components N is small by comparison with the size n of a no-fail sample (thus meeting the requirements that $n \gg N$ and $Nf \ll 1$), the component reliability corresponding to a confidence of 75.7% could be used as a true (unqualified) pass rate for calculation of the reliability of the system.

The degree of approximation involved in using component reliability corresponding to a 70% confidence (as a compromise between the 63.2% appropriate for series systems and the 75.7% appropriate for parallel systems) is shown in Table 3.

Appendix B

Sifting Effect of Small Sample No-Fail Tests

No-fail tests of n samples can be thought of as a sieve, separating the better lots from the poorer lots. Suppose the lots come in boxes of 100 parts per box, and that a warehouse contains 101 million boxes. Suppose further, that all pass rates are equally likely. Then, there are a million boxes containing no good parts, a million with one, a million with 2, . . . , a million with 99 good parts, and a million boxes with all 100 parts good. If we knew how many good parts were in each box, we could stack all those with each number of good parts in a separate pile. If these piles were arranged in order of the number of good parts, the tops of the piles would look like the upper line ($n = 0$) in Fig. B-1.

Suppose that one part is taken out of each box and tested. Some of the tests will reveal defective parts. When this happens, we will remove that box from the stack. The probability that we will not remove a box from its pile, when all the boxes in that pile have p good parts (and $100 - p$ defective parts), is simply $p/100$. Since there were a million boxes in the pile there will now be $(p/100) \times 10^6$ boxes left. The tops of the piles would look like the line marked $n = 1$ in Fig. B-1.

Similarly, if we take a second sample from each box, and again throw out the boxes that yielded a defective sample, the tops of the piles would look like the line marked $n = 2$. There were $(p/100) \times 10^6$ boxes in the p th pile. Only $p/100$ (approximately) of these boxes passed the second test, leaving $(p/100)^2 \times 10^6$ boxes in the p th pile. Figure B-1 shows the appearance of the top of the stack after several siftings. The number of boxes left in the p th pile, after any box yielding a defective sample in n samples has been discarded, is $(p/100)^n \times 10^6$.

What has all this to do with reliability and confidence? The sorting in the previous paragraphs can only be done conceptually, of course, since the number of good parts in each lot (box) is not known to us. Consequently, accepting a lot after it passes an n -sample no-fail test is like taking all the piles that are left, mixing up the boxes, and drawing one at random. How many good parts does it contain? What is the chance that a system composed of parts from this box will work?

The probability that a part taken from any particular box will work is simply the number of good parts in that

box, p , divided by 100. But, by definition, this is the reliability of the parts in that box. On drawing a box, however, we do not know what this reliability is. We do know how many boxes there were with each reliability. Consequently, we can state a probability that the reliability of the drawn box is at least R , for any value of R between 0 and 1. Again, we note that this probability is, by definition, the confidence, γ , that the reliability is R . The relationship between reliability and confidence resulting from an n -sample, no-fail sifting sequence applied to an initial lot distribution in which all pass rates were equally likely (usually a conservative assumption) can be easily expressed by dividing the number of boxes with an equal or higher reliability than R by the total number of boxes left. Thus,

$$\gamma = \frac{\int_R^1 (p/100)^n \times 10^6 dp}{\int_0^1 (p/100)^n \times 10^6 dp} = 1 - R^{n+1}$$

The quantity which we have called inherent reliability is the total fraction of good parts in the boxes remaining. Since there are p good parts in each box in the p th pile, we have

$$R_i = \frac{\int_0^1 (p/100)^n \times 10^6 \times p dp}{\int_0^1 (p/100)^n \times 10^6 dp} = \frac{n+1}{n+2}$$

Components, however, are only of interest because we wish to combine them to make systems. Immediately, we are faced with finding those (γ, R) pairs which should be chosen to give specified system pass rates. If we continue with the assumption that the n -sample, no-fail test has been used to preferentially sift an initial population in which all pass rates were equally likely, we seek that confidence level, γ , whose associated reliability, R_γ , will give the same average system pass rate as if R_γ were the true reliability of the lot used.

It was shown in Appendix A that the result of this quest depends on the logical connection of components in the particular system being studied, but that the appropriate

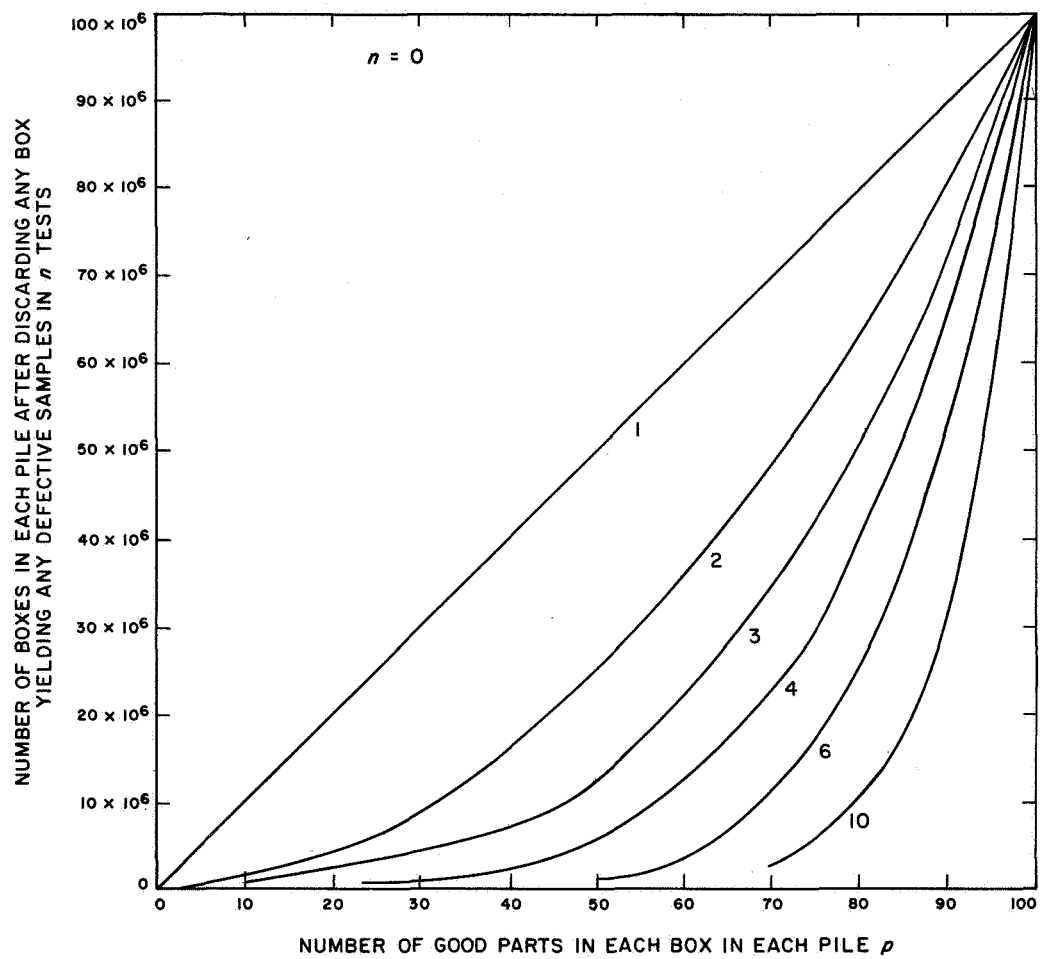


Fig. B-1. Number of lots of each reliability after n no-fail tests

confidence level ranges from 63% to 76%. The suggestion was made in Section V that component reliabilities corresponding to a confidence level of 70% may give a satisfactory approximation for most uses.

We have seen in Fig. B-1 that low pass rates are indeed rapidly screened out by the small sample no-fail tests. Suppose, however, we are concerned with an application in which we would consider a 99% reliability to be low. Are small sample no-fail tests still useful? By considering the fact that 98 samples must be drawn to raise the inher-

ent reliability to 99%, we see that large samples (i.e., at least 98) will be required.

To fully answer this question, consider Fig. B-2, which shows the sifting effect on lots of high-reliability components. This figure shows that relatively large samples must still be tested to obtain an appreciable sifting effect even if only lots with pass rates exceeding 99% were initially present. The implication is that testing by attributes (go, no-go) is not efficient for demonstration or determination of high reliabilities.

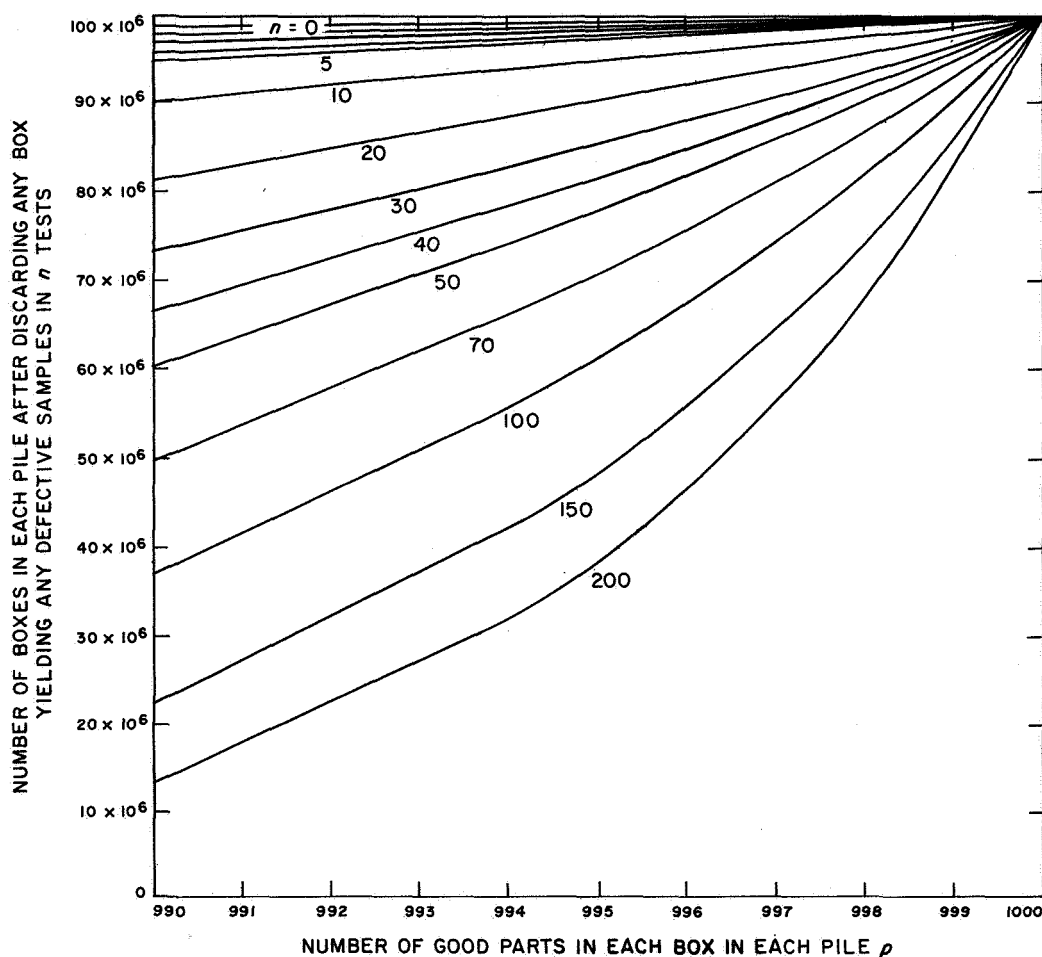


Fig. B-2. Sifting effect on high-reliability lots